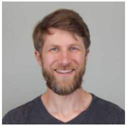
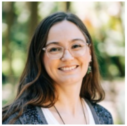
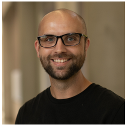
Locally Equivalent Weights for Bayesian MrP

Ryan Giordano, Alice Cima, Erin Hartman, Jared Murray, Avi Feller

UT Austin Statistics Seminar
September 2025




What are we weighting for?[1]

$$\text{Target average response} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j \approx \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i = \text{Weighted survey average response}$$

We can't check this, because we don't observe y_j.

[1] Pun attributable to Solon, Haider, and Wooldridge (2015)

What are we weighting for?[1]

$$\text{Target average response} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j \approx \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i = \text{Weighted survey average response}$$

We can't check this, because we don't observe y_j. But we can check whether:

$$\frac{1}{N_T}\sum_{j=1}^{N_T} \mathbf{x}_j = \frac{1}{N_S}\sum_{i=1}^{N_S} w_i \mathbf{x}_i$$

Such weights satisfy "covariate balance" for $\mathbf{x}$.

You can check covariate balance for any calibration weighting estimator, and any function $f(\mathbf{x})$.

[1] Pun attributable to Solon, Haider, and Wooldridge (2015)

What are we weighting for?[1]

$$\text{Target average response} = \frac{1}{N_T}\sum_{j=1}^{N_T} y_j \approx \frac{1}{N_S}\sum_{i=1}^{N_S} w_i y_i = \text{Weighted survey average response}$$

We can't check this, because we don't observe y_j. But we can check whether:

$$\frac{1}{N_T}\sum_{j=1}^{N_T} \mathbf{x}_j = \frac{1}{N_S}\sum_{i=1}^{N_S} w_i \mathbf{x}_i$$

Such weights satisfy "covariate balance" for $\mathbf{x}$.

You can check covariate balance for any calibration weighting estimator, and any function $f(\mathbf{x})$.

Even more, covariate balance is the criterion for a popular class of calibration weight estimators:

Raking calibration weights

"Raking" selects weights that

- Are as "close as possible" to some reference weights
- Under the constraint that they balance some selected regressors.

[1] Pun attributable to Solon, Haider, and Wooldridge (2015)

Balance checks as sensitivity analysis

One reason to balance $f(\mathbf{x})$ is because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance checks as sensitivity analysis

One reason to balance $f(\mathbf{x})$ is because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance–informed sensitivity check (BISC) (informal)

Pick a small $\delta > 0$ and an $f(\cdot)$. Define a *new response variable* $\tilde{y}$ such that

$$\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x}).$$

We know the change this is supposed to induce in the target population.

Covariate balance checks whether our estimators produce the same change.

Balance checks as sensitivity analysis

One reason to balance $f(\mathbf{x})$ is because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance–informed sensitivity check (BISC) (formal)

Pick a small $\delta > 0$ and an $f(\cdot)$. Define a *new response variable* $\tilde{y}$ such that

$$\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x}).$$

We know the expected change this perturbation produces in the target distribution:

$$\mathbb{E}\left[\mu(\tilde{y}) - \mu(y)|\mathbf{x}\right] = \frac{1}{N_T}\sum_{j=1}^{N_T}\left(\mathbb{E}\left[\tilde{y}|\mathbf{x}_p\right] - \mathbb{E}\left[y|\mathbf{x}_p\right]\right) = \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j)$$

Then, check whether your estimator $\hat{\mu}(\cdot)$ produces the same change for observed $\tilde{y}, y$:

$$\underbrace{\hat{\mu}(\tilde{y}) - \hat{\mu}(y)}_{\substack{\text{Replace weighted averages} \\ \text{with changes in an estimator}}} \overset{\text{check}}{\approx} \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j).$$

Balance checks as sensitivity analysis

One reason to balance $f(\mathbf{x})$ is because we think $\mathbb{E}\left[y|\mathbf{x}\right]$ might plausibly vary $\propto f(\mathbf{x})$, and want to check whether our estimator can capture this variability.

Balance–informed sensitivity check (BISC) (formal)

Pick a small $\delta > 0$ and an $f(\cdot)$. Define a *new response variable* $\tilde{y}$ such that

$$\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x}).$$

We know the expected change this perturbation produces in the target distribution:

$$\mathbb{E}\left[\mu(\tilde{y}) - \mu(y)|\mathbf{x}\right] = \frac{1}{N_T}\sum_{j=1}^{N_T}\left(\mathbb{E}\left[\tilde{y}|\mathbf{x}_p\right] - \mathbb{E}\left[y|\mathbf{x}_p\right]\right) = \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j)$$

Then, check whether your estimator $\hat{\mu}(\cdot)$ produces the same change for observed $\tilde{y}, y$:

$$\underbrace{\hat{\mu}(\tilde{y}) - \hat{\mu}(y)}_{\substack{\text{Replace weighted averages} \\ \text{with changes in an estimator}}} \overset{\textbf{check}}{\approx} \delta\frac{1}{N_T}\sum_{j=1}^{N_T} f(\mathbf{x}_j).$$

When $\hat{\mu}(\cdot) = \hat{\mu}_{\mathrm{CW}}(\cdot)$, BISC recovers the standard covariate balance check.

We will use $\hat{\mu}(\cdot) = \hat{\mu}_{\mathrm{MrP}}(\cdot)$.

BISC for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

BISC for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

Problem: $\hat{\mu}_{\text{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$ may be small.

BISC for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$.

Problem: $\hat{\mu}_{\text{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\text{MrP}}(\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\text{MrP}} := \frac{d}{dy_i}\hat{\mu}_{\text{MrP}}(y).$$

BISC for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$.

Problem: $\hat{\mu}_{\mathrm{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\mathrm{MrP}} := \frac{d}{dy_i} \hat{\mu}_{\mathrm{MrP}}(y).$$

Computation: The weights are given by weighted averages of posterior covariances[2].

They can be easily computed with standard software[3] **without re–running MCMC**.

[2] **G.**, Broderick, and Jordan 2018.
[3] We use `brms` (Bürkner 2017).

BISC for MrP

Suppose I have $\tilde{y}$ such that $\mathbb{E}\left[\tilde{y}|\mathbf{x}\right] = \mathbb{E}\left[y|\mathbf{x}\right] + \delta f(\mathbf{x})$.
Now I need to evaluate $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$.

Problem: $\hat{\mu}_{\mathrm{MrP}}(\cdot)$ is computed with MCMC.

- Each MCMC run typically takes hours, and
- Output is noisy, and $\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y)$ may be small.

MrP Local Equivalent Weights (MrPlew)

Form the first-order Taylor series approximation

$$\hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) \approx \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} (\tilde{y}_i - y_i) \quad \text{where} \quad w_i^{\mathrm{MrP}} := \frac{d}{dy_i} \hat{\mu}_{\mathrm{MrP}}(y).$$

Use in BISC: For a wide set of judiciously chosen $f(\cdot)$, check

$$\delta \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} f(\mathbf{x}_i) \overset{\text{check}}{\approx} \delta \frac{1}{N_T} \sum_{j=1}^{N_T} f(\mathbf{x}_j).$$

Theory

BISC Theorem: (sketch)

We state conditions for Bayesian hierarchical logistic regression under which

$$\left|\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y) - \delta \sum_{i=1}^{N_S} w_i^{\text{MrP}} f(\mathbf{x}_i)\right| = \text{Small?}$$

[2] Donsker with uniformly bounded $\mathbb{E}\left[\mathbf{x} f(\mathbf{x})\right]$.

[3] **G.** and Broderick 2024; Kasprzak, **G.**, and Broderick 2025.

BISC Theorem: (sketch)

We state conditions for Bayesian hierarchical logistic regression under which

$$\left| \hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) - \delta \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} f(\mathbf{x}_i) \right| = O(\delta^2)$$

[2] Donsker with uniformly bounded $\mathbb{E}\left[\mathbf{x} f(\mathbf{x})\right]$.

[3] **G.** and Broderick 2024; Kasprzak, **G.**, and Broderick 2025.

BISC Theorem: (sketch)

We state conditions for Bayesian hierarchical logistic regression under which

$$\left|\hat{\mu}_{\text{MrP}}(\tilde{y}) - \hat{\mu}_{\text{MrP}}(y) - \delta \sum_{i=1}^{N_S} w_i^{\text{MrP}} f(\mathbf{x}_i)\right| = O(\delta^2) \text{ as } N \to \infty$$

---

[2] Donsker with uniformly bounded $\mathbb{E}\left[\mathbf{x} f(\mathbf{x})\right]$.

[3] **G.** and Broderick 2024; Kasprzak, **G.**, and Broderick 2025.

BISC Theorem: (sketch)

We state conditions for Bayesian hierarchical logistic regression under which

$$\sup_{f \in \mathcal{F}} \left| \hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) - \delta \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} f(\mathbf{x}_i) \right| = O(\delta^2) \text{ as } N \to \infty$$

For a very broad class[2] of $\mathcal{F}$.

Uniformity justifies searching for "imabalanced" f.

[2] Donsker with uniformly bounded $\mathbb{E}\left[\mathbf{x} f(\mathbf{x})\right]$.

[3] **G.** and Broderick 2024; Kasprzak, **G.**, and Broderick 2025.

BISC Theorem: (sketch)

We state conditions for Bayesian hierarchical logistic regression under which

$$\sup_{f \in \mathcal{F}} \left| \hat{\mu}_{\mathrm{MrP}}(\tilde{y}) - \hat{\mu}_{\mathrm{MrP}}(y) - \delta \sum_{i=1}^{N_S} w_i^{\mathrm{MrP}} f(\mathbf{x}_i) \right| = O(\delta^2) \text{ as } N \to \infty$$

For a very broad class[2] of $\mathcal{F}$.

Uniformity justifies searching for "imabalanced" f.

The uniformity result builds on our earlier work on uniform and finite–sample error bounds for Bernstein–von Mises theorem–like results[3].

[2] Donsker with uniformly bounded $\mathbb{E}\left[\mathbf{x} f(\mathbf{x})\right]$.

[3] **G.** and Broderick 2024; Kasprzak, **G.**, and Broderick 2025.

Future work

Note that there was no talk of correct specification for the data you have.

That was a foregone conclusion when we started looking at equivalent weights!

How do you peform model checking with sensitivity analysis?

Existing methods evaluate whether the analysis changes “a lot” when you:

- Parametrically perturb the model (e.g. fit a richer model class)
- Non–parameterically perturb the data (e.g. produce gross outliers)

The problem is:

- How much is “a lot”?
- Non–parametric data perturbations are hard to reason about
- It’s hard to say whether parametric model changes are enough

Instead, we

- Parametrically perturb the data
- Observe whether our model could detect the change

- Know exactly the expected change (don’t have to decide on what “a lot” means)
- Easy to reason about whether the data perturbation is reasonable
- Don’t need to propose an alternative model, instead study the model you have

Related and future work

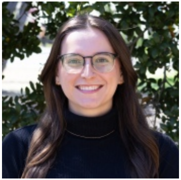
Student contributions and future work:

- **Alice Cima** contributed significantly to this work
- **Vladimir Palmin** is working on extending MrPlew to `lme4`
- **Sequoia Andrade** is working on generalizing to other local sensitivity checks
- **Lucas Schwengber** is working on novel flow–based techniques for local sensitivity

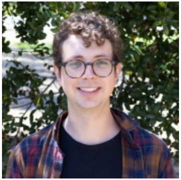
Alice Cima

No picture!

Vladimir Palmin



Sequoia Andrade



Lucas Schwengber

References

Bürkner, Paul-Christian (2017). "brms: An R Package for Bayesian Multilevel Models Using Stan". In: *Journal of Statistical Software* 80.1, pp. 1–28. DOI: 10.18637/jss.v080.i01.

G. and T. Broderick (2024). *The Bayesian Infinitesimal Jackknife for Variance*. arXiv: 2305.06466 [stat.ME]. URL: https://arxiv.org/abs/2305.06466.

G., T. Broderick, and M. I. Jordan (2018). "Covariances, robustness and variational bayes". In: *Journal of machine learning research* 19.51.

Kasprzak, M., **G.**, and T. Broderick (2025). *How good is your Laplace approximation of the Bayesian posterior? Finite-sample computable error bounds for a variety of useful divergences*. arXiv: 2209.14992 [math.ST]. URL: https://arxiv.org/abs/2209.14992.

Solon, G., S. Haider, and J. Wooldridge (2015). "What are we weighting for?" In: *Journal of Human resources* 50.2, pp. 301–316.